基于 BERTopic 的国内智慧图书馆“科学—技术”主题研究*

高道斌
（大连理工大学科学学与科技管理研究所暨 WISE 实验室，大连　116024）

摘　要：[**目的 / 意义**] 识别和分析国内智慧图书馆“科学—技术”主题，为该领域科学研究与技术开发提供情报支撑，为科学和技术的知识双向流动提供参考。[**方法 / 过程**] 以 CNKI 期刊数据库 5436 篇中文论文与 incoPat 专利数据库 1416 件中国专利为数据源，分别表征科学载体与技术载体，基于 BERTopic 模型进行“科学—技术”主题识别，分析主题间聚类结果、重要研究主题与主题演化趋势。在此基础上，分析各重要科学研究主题与技术研究主题间的交互关系，提出未来相关研究启示。[**结果 / 结论**] 基于 BERTopic 模型将科学研究领域划分为 4 类，其重要科学主题囊括 5 项；将技术研究领域划分为 3 类，其重要技术主题包括 3 项；识别“科学—技术”交互关系共 7 项。就此提出遵循政策指引与场景化应用导向，促进研究成果从理论向实践转化；拓展面向公共智慧图书馆的服务研究，实现“公共馆—高校馆”并重；提升智慧图书馆各类主题研究的可持续性，夯实科学、技术探索深度；注重引入技术信息支撑，增加“科学—技术”交互程度共 4 点研究启示。

关键词：智慧图书馆　BERTopic　科学主题　技术主题　交互现状
分类号：G250.76
DOI：10.31193/SSAP.J.ISSN.2096-6695.2025.01.14

0　引　言

伴随数智技术的向深发展与用户信息服务需求方向的转变，智慧化转型成为我国图书馆发展的必然趋势。在此情境下，面向智慧图书馆的研究已成为当前国内图书情报领域热点。同时，在政策规划方面，《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》[1] 将智慧图书馆建设作为数字化转型重要内容。智慧图书馆的发展与建设，既需要技术作为基础架构，为各类智慧化服务，如图书归还、标签自动整理、环境感知等提供技术赋能，也需要科学研究为其指引路径与方向，如发现问题、提出理论框架、规划发展路径等。

* 本文系教育部哲学社会科学研究重大课题攻关项目“基础研究领域颠覆性科研成果识别与我国基础研究能力提升研究”（项目编号：22JZD021）的研究成果之一。
[作者简介] 高道斌，男，博士生，研究方向为科学计量学、知识图谱，Email：gaodaobin1998@163.com。

当前，国内面向智慧图书馆的发展现状分析主要以面向科学论文的研究为主。如王飞等[2]以 2021 年前国内发布的智慧图书馆论文为数据源，采用词频分析、社会网络分析等方法，对智慧图书馆文化研究现状进行量化分析，揭示了研究热点。丁安等[3]以中国知网为数据源，对 2009—2018 年间我国智慧图书馆研究现状和热点内容进行分析，揭示智慧图书馆服务模式、现存问题和应对策略等。李显志等[4]以维普网、中国知网、万方数据库为数据源进行文献计量分析，发现我国学者对智慧图书馆基本问题进行了深入探讨，初步构想了智慧图书馆建设和服务模式。此外，少数研究以专利为数据源对国内智慧图书馆技术发展现状进行剖析。如裴宏娇等[5]以国家知识产权局数据库为数据源，对 2013—2022 年国内智慧图书馆专利进行基础计量分析，发现 RFID、人工智能等技术在我国智慧图书馆的应用尚处于探索发展时期，提出未来可以通过不断加强主体合作、利用科研院所的专利技术、夯实理论基础等方式提高技术创新能力。赵昆[6]以 2022 年前的国内智慧图书馆领域专利为分析载体，利用 CiteSpace 和 VOSviewer 对专利关键词和专利权人关系进行共现分析，等等。

可见，国内智慧图书馆发展现状研究侧重从单一数据载体出发，如论文、专利等，去识别领域研究趋势、热点等，鲜有研究兼顾两者的分析，更缺乏面向科学（论文）与技术（专利）间交互关系的探讨。当前，相关研究多采用简单文献或专利计量方法，如词频统计、社会网络分析、数量变化趋势研判等，较少研究基于深度学习方法实现对领域内研究主题的识别与分析，导致对研究现状、研究主题等的揭示粒度较粗，难以满足细粒度文本语义分析需求。基于深度学习的 BERTopic 主题建模方法能够较好解决上述问题。该模型是基于 BERT 模型无监督深度学习主题建模工具，具有强大语言特征提取能力，考虑了文本词汇背后的语义、结构和序列，使得到的主题可解释性较强并在一定程度上实现词义消歧[7]。同时，该模型无需人为设置主题数量，相较 LDA、DTM 等文本主题模型具有一定优势。目前，BERTopic 主题模型已被国内学者应用于数字政府治理领域的主题识别与内容分析[8]、老年人健康信息需求主题演化研究[9]、国外信息资源管理研究进展分析[10]等，证明了其具备良好的文本挖掘与主题识别能力，普适性较好。

综上，为解决国内智慧图书馆研究缺乏“科学—技术”交互分析、研究方法对文本语义深入挖掘能力缺乏、揭示主题信息粒度较粗等问题，本文基于 BERTopic 主题模型，以论文表征科学载体、以专利表征技术载体，从“科学—技术”双重视角出发，识别智慧图书馆“科学—技术”主题，分析“科学—技术”交互现状，以期为该领域科学研究与技术开发提供情报支撑，为“科学—技术”知识双向流动提供数据参考。

1 研究流程

本文研究流程主要包括数据获取与处理、BERTopic 模型应用、识别与分析 3 个部分。

（1）数据获取与处理：分别以论文、专利表征科学载体、技术载体，在分别剔除论文选题指南和荐读等文献，以及摘要缺失专利后，对科学载体、技术载体集合的摘要进行文本语料预处理与清洗，形成原始语料集合。

（2）BERTopic 模型应用：①将经过预处理与清洗后的摘要语料嵌入基于 BERT 的语言预

处理模型，利用 Transformer 结构计算摘要词向量；②通过 UMAP 对上一步得到的句嵌入进行降维，提高聚类算法效率和运算精度；③利用 HDBSCAN 算法聚合相似属性的文献数据，自动选择主题聚类数量，形成主题簇；④提取各科学、技术主题下 c-TF-IDF 权重排名前列主题关键词，并利用层次聚类功能进行主题层次聚类可视化展现。

（3）识别与分析：①分别以科学载体与技术载体的历年数量分布特征为依据，分析两类载体数量变化趋势；②对各科学、技术研究主题的层次聚类结果、重要研究主题（以主题下文档分布数量表征主题强度，对大于等于平均主题强度的科学、技术研究主题进行二次人工归类与命名后的主题结果）及重要研究主题演化趋势进行深入分析；③基于各重要科学、技术研究主题间，所有文档集合摘要语料的平均余弦相似度，对“科学—技术”主题交互关系进行分析。上述研究流程见图 1。

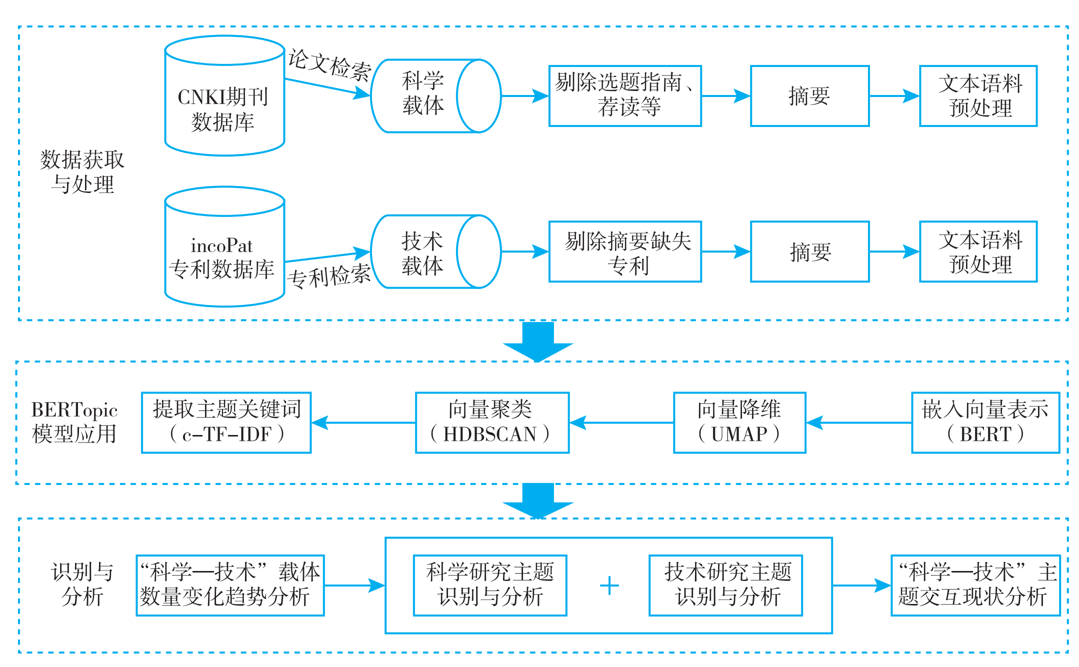


图 1 研究流程

2 国内智慧图书馆“科学—技术”主题识别与分析

2.1 数据获取

科学载体以 CNKI 期刊数据库为数据源，参考前期研究成果[11]，并对其改进与扩充，最终确定检索式为：“SU%=（智慧 OR 智能 OR 自动 OR 感知 OR 云计算 OR 大数据 OR 物联网）AND SU%=（图书馆）”。由于“智慧图书馆”这一概念于 2003 年正式提出[12]，将发文日期设置为 2003 年 1 月 1 日至 2024 年 8 月 1 日，检索日期为 2024 年 8 月 1 日，在检索后进一步剔除期刊选题指南、书目推荐等类型文献，获取国内智慧图书馆领域论文共 5436 篇，保存上述论文的题目、作者、摘要、发表时间等字段信息，作为此次研究科学载体数据源。

技术载体以 incoPat 专利数据库为数据源，检索式同于科学载体检索式，为适应数据库检索需求，仅修改检索字段标记：“TIAB=（智慧 OR 智能 OR 自动 OR 感知 OR 云计算 OR 大数据 OR 物联网）AND TIAB=（图书馆）”，将申请日期设置为 2003 年 1 月 1 日至 2024 年 8 月 1 日，专利审查局设置为中国，专利类型设置为授权发明与发明申请，检索日期为 2024 年 8 月 1 日，在检索后进一步剔除摘要缺失专利，获取国内智慧图书馆领域专利共 1416 件，保存上述专利的标题、公开号、摘要、申请时间等字段信息，作为此次研究技术载体数据源。

2.2 “科学—技术”载体数量变化趋势分析

在 2003—2024 年时间范围内，分别将科学载体与技术载体的分布数量作为数据源，绘制“科学—技术”载体数量整体变化趋势图，从“量”与“势”两个角度揭示领域内发展概况，结果见图 2。

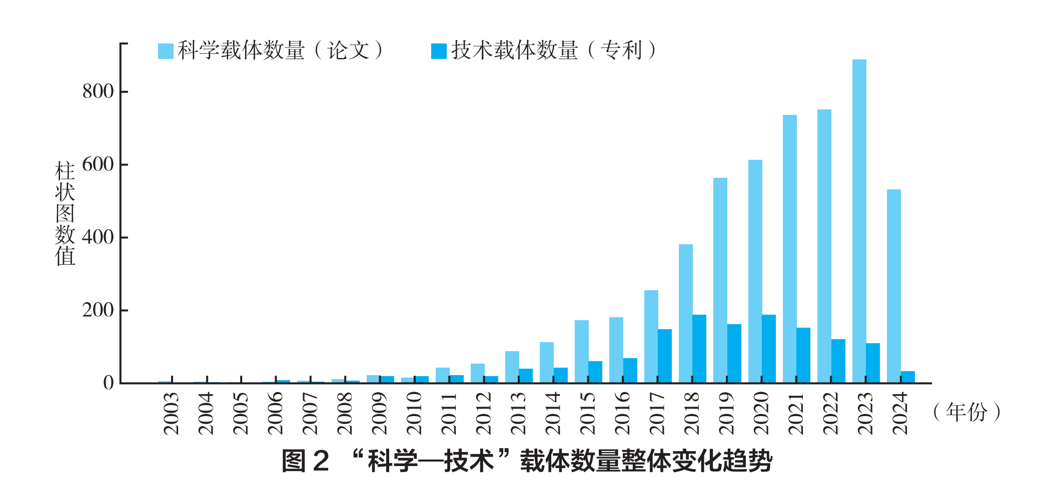


图 2 “科学—技术”载体数量整体变化趋势

如图 2 所示，在 2003—2010 年间，科学载体与技术载体的数量大致相当。在 2011—2024 年间，科学载体历年分布数量均高于技术载体。单从数量分布角度可以判断，早期科学研究与技术研究并驾齐驱，但在 2011 年后科学研究活跃程度高于技术研究。从数量分布趋势角度来看，两者在 2010 年前载体数量基数与变化幅度均较小，科学载体在 2011—2023 年间分布数量相较 2003—2010 年间呈显著上升趋势，技术载体数量在 2011—2018 年间整体呈上升趋势，但在 2020—2024 年间整体呈显著下降趋势。

2.3 科学研究主题识别与分析

2.3.1 科学研究主题层次聚类结果分析

基于 BERTopic 模型识别国内智慧图书馆领域科学研究主题共 39 个，通过 BERTopic 模型层次聚类功能，展现所有科学研究主题的主题名称与层次聚类情况，见图 3。

如图 3 所示，通过层次聚类图谱颜色分布情况可知，国内智慧图书馆科学研究领域大致可以划分为 4 种类别。

（1）深绿色聚类结果：智慧图书馆数据治理与社会利他行为。该聚类结果下的科学研究主题包括：主题 36（治理 _ 数据 _ 圈层）、主题 37（社会 _ 利他行为 _ 信息）和主题 27（精神 _ 古代 _ 图书馆学）。王静等[13]指出高校图书馆智慧服务数据治理是智慧图书馆重要研究领域，数字

孪生技术的出现为数据治理过程中的信息物理融合提供了新的解决途径。在数据时代，数据治理和智慧服务成为图书馆发展关注焦点之一[14]。在馆员视角下，利他行为是图书馆全心全意服务读者、不图回报满足读者的行为。图书馆的利他行为体现在馆建、管理、服务和发展等方面，图书馆的利他行为展示了图书馆人高度的社会责任和社会义务，显示了图书馆人高尚道德以及谦逊修养[15]。

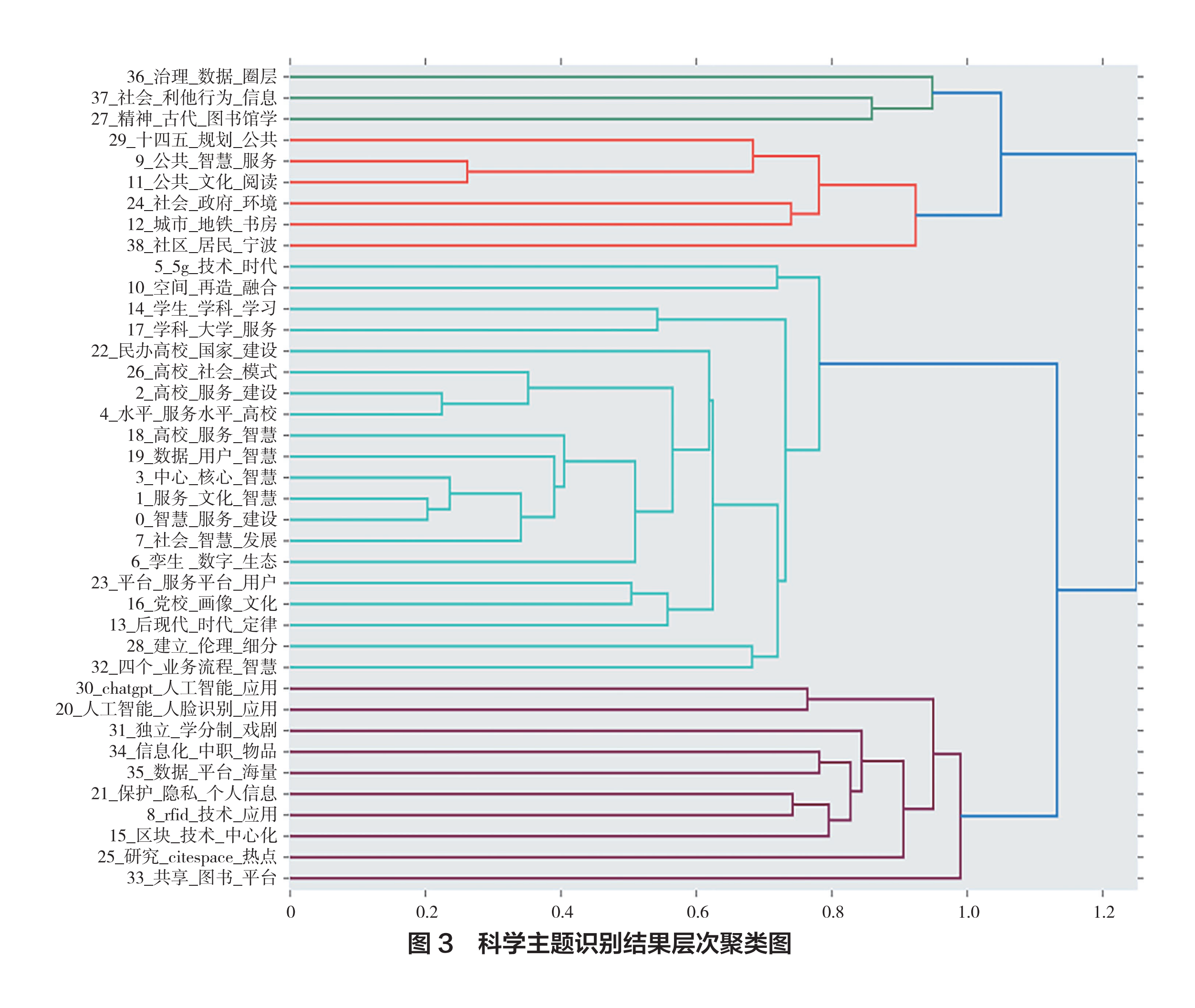


图 3 科学主题识别结果层次聚类图

（2）红色聚类结果：十四五规划下的智慧公共文化服务。该聚类结果下的科学研究主题包括：主题 29(十四五 _ 规划 _ 公共)、主题 9(公共 _ 智慧 _ 服务) 和主题 11(公共 _ 文化 _ 阅读) 等。进入“十四五”时期，开启了全面建设社会主义现代化国家新征程，图书馆转型发展迎来重要机遇时期。《“十四五”公共服务规划》[16] 提出要“推进全国智慧图书馆体系建设……等设施按规定免费或优惠开放”。“十四五”时期，大数据、云计算、人工智能等智慧技术持续深入发展和应用，公共图书馆智慧服务方式变革迎来新契机[17]。杜希林等[18] 的研究表明数智赋能公共图书馆服务体系建设存在优势与风险，需要从用户、馆员、资源、空间、技术 5 个路径寻求赋能切入点及方式，推动公共图书馆服务体系创新。秦春阳[19] 在深入分析公共图书馆智慧服务体系内部结构要素基础上，聚焦智慧服务体系的优化发展，提出更新服务观念以引领智慧服务体系发展方向，革新服务手段以促进智慧服务体系质量变革等观点。

（3）青绿色聚类结果：高校智慧图书馆服务平台建设与空间再造。该聚类结果下的科学研究主题包括：主题 2（高校 _ 服务 _ 建设）、主题 18（高校 _ 服务 _ 智慧）和主题 23（平台 _ 服务平台 _ 用户）等。在图书馆智慧化变革与发展大潮下，国内高校图书馆积极探索从资源管理模式到服务管理模式的转型[20]。杜芃诺[21]从 3 个方面提出构建高校智慧图书馆服务平台，分别是推进精准服务机制，打造智慧人文标杆；构建智慧阅读资源体系，贯通资源利用渠道；虚实结合强化阅读体验，技术驱动扩展空间服务。马捷等[22]结合高校图书馆服务对象及业务特点，从智慧服务、智慧建筑、智慧管理 3 个维度对高校智慧图书馆功能结构进行整体设计和构建。车宝晶等[23]通过调研 42 家“双一流”高校图书馆，总结数智赋能学科服务现状与现实困境，提出数智赋能高校图书馆学科服务转型路径，包括利用数智技术丰富智慧学科服务场景，构建虚实交融、功能优化的智慧学科服务空间等。

（4）酒红色聚类结果：智慧图书馆技术赋能与用户隐私管理。该聚类结果下的科学研究主题包括：主题 30（chatgpt_ 人工智能 _ 应用）、主题 8（rfid_ 技术 _ 应用）和主题 21（保护 _ 隐私 _ 个人信息）等。人工智能、大数据、云计算等技术让图书馆数据的价值得以转换，然而数据价值发掘过程必然涉及读者隐私[24]。图书馆在积极寻求技术赋能的同时要保持审慎态度，在充分尊重用户隐私和保障用户信息安全基础上，发展适用于图书馆需求的智慧化服务创新，规避技术背后的安全风险[25]。郭军[26]在分析智慧环境下图书馆用户个人信息受到侵害原因基础上，构建包含预防预警、过程控制、问责追究、制度监管 4 个模块的图书馆用户个人信息保护工作框架。梅振荣等[27]提出将区块链技术应用到智慧图书馆用户的隐私保护，借鉴时间戳、哈希函数、Merkle 可信树及共识机制等技术，构建集智慧链的防篡改、隐私加密和安全存储机制于一体的用户隐私保护架构模型。

2.3.2 重要科学研究主题分析

国内智慧图书馆领域主题强度（以主题下论文分布数量表示）大于等于均值的科学研究主题共计 11 项。进一步结合各主题间关键词语义进行二次人工归类与命名，最终获得重要科学研究主题 5 项，归类前后主题名称、主题强度情况见表 1。

表 1　国内智慧图书馆领域重要科学研究主题分布情况

归类后重要科学研究主题名称	归类后主题强度（论文数量）	原始科学研究主题名称	主题强度（论文数量）
智慧图书馆智慧文化服务	1293	主题 0- 智慧 _ 服务 _ 建设	683
		主题 1- 服务 _ 文化 _ 智慧	610
高校智慧图书馆服务水平	720	主题 2- 高校 _ 服务 _ 建设	420
		主题 4- 水平 _ 服务水平 _ 高校	300
智慧图书馆新兴技术应用	444	主题 5-5g_ 技术 _ 时代	153
		主题 6- 孪生 _ 数字 _ 生态	148
		主题 8-rfid_ 技术 _ 应用	143
智慧图书馆学习中心建设	425	主题 3- 中心 _ 核心 _ 智慧	315
		主题 10- 空间 _ 再造 _ 融合	110
公共智慧图书馆文化服务体系	272	主题 7- 社会 _ 智慧 _ 发展	143
		主题 9- 公共 _ 智慧 _ 服务	129

如表 1 所示,“智慧图书馆智慧文化服务”主题强度最高，表示该研究主题在国内智慧图书馆相关研究中处于核心位置，引领作用突出。该主题关注通过智慧图书馆为用户提供各类智慧文化服务，以增强用户体验，反映智慧图书馆发展中对文化维度的实践探讨是重要研究议题。智慧图书馆赋予文化服务智慧化与主动性，提升智慧文化服务水平、提供场景化智慧文化服务及化解各类智慧文化服务中存在的服务问题，为该主题下重要研究内容。

“高校智慧图书馆服务水平”主题强度排名第 2，表明国内智慧图书馆领域面向如何提升高校图书馆智慧化服务水平、促进学习资源高效利用及满足师生多元化学习需求等方面展开了广泛探讨。高校智慧图书馆是相较于公共智慧图书馆等的重要智慧化主体之一，不仅面向高校图书馆丰富馆藏资源展开智慧挖掘、分析、推送、检索等，为高校师生提供高效、个性化知识获取途径，还注重面向知识产权管理、科技查新等服务内容的优化建设，更加聚焦面向校内科研资源的智慧化。

“智慧图书馆新兴技术应用”主题强度排名第 3，体现了各类技术要素对国内智慧图书馆研究的重要支撑作用。智慧图书馆智慧化需要借助各类技术实现，如通过 RFID、物联网技术、云计算技术等实现图书馆内图书标签自动识别、归类、定位等；又如大语言模型、增强现实、数字孪生等新兴技术赋能智慧图书馆实现智慧用户交流、沉浸式阅读等服务功能。新兴技术应用使智慧图书馆不断突破服务界限，提升智慧化水平。

“智慧图书馆学习中心建设”主题强度排名第 4，说明本领域研究者逐渐将对智慧图书馆主体研究分散至对各类场景式智慧图书馆主体研究。以智慧图书馆未来学习中心为例，通过借助各类数智技术对图书馆学习空间进行改造，实现面向各类学习者的学习资源推送、学习画像构建、学习环境优化等服务，为当下热点研究议题。

“公共智慧图书馆文化服务体系”主题强度排名第 5，表明本领域研究者对该议题重视。相较于高校智慧图书馆，公共智慧图书馆服务群体更加广泛，并且服务内容更加分散，基于各类服务群体知识结构差异性，公共智慧图书馆整合各类文化资源，利用数智化技术手段，构建全方位、多层次、高效率文化服务体系，为各类服务群体提供优质阅读服务。

2.3.3 重要科学研究主题演化趋势分析

以时间为横轴、重要科学研究主题的主题强度为竖轴，面向国内智慧图书馆领域 5 项重要科学研究主题进行演化趋势分析，结果见图 4。

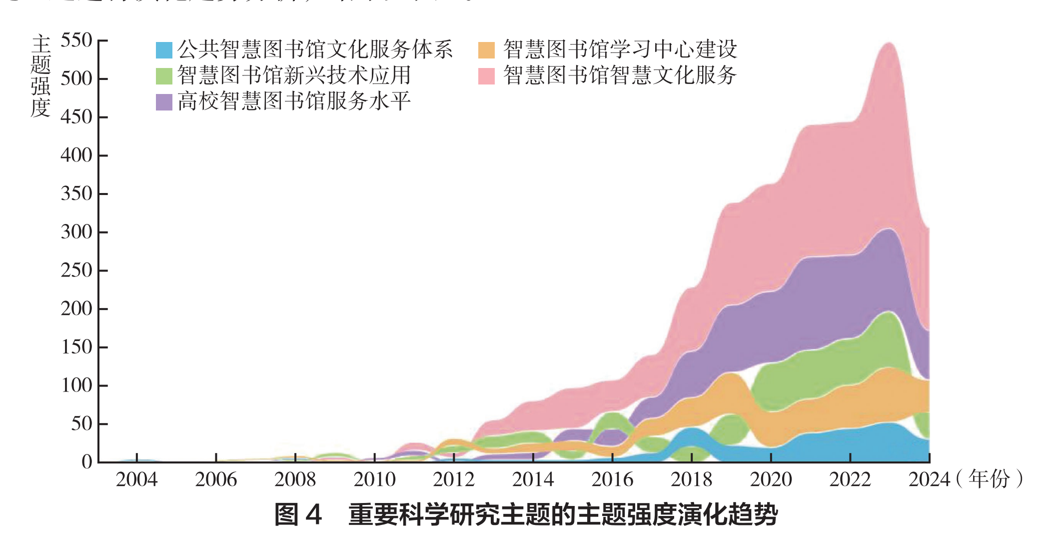


图 4 重要科学研究主题的主题强度演化趋势

如图 4 所示，在 2012 年前，各重要科学研究主题的主题强度均较低，在 2012 年后，各重要科学研究主题的主题强度显著增加，且整体保持稳定上涨趋势。其中，“智慧图书馆智慧文化服务”主题在 2012 年后时间段内，主题强度变化趋势较大，具体表现为在 2012—2023 年整体呈稳定上升趋势，且在 2018—2023 年间显著上升，说明该研究主题研究热度持续升温，为领域内重要研究议题。其余 4 项重要科学研究主题在 2012—2018 年间主题强度差距较小，但在 2018 年后，“高校智慧图书馆服务水平”主题强度同其他 3 项差距显著，表明领域内学者逐渐聚焦面向高校智慧图书馆服务有关研究。此外，“智慧图书馆学习中心建设”“智慧图书馆新兴技术应用”“公共智慧图书馆文化服务体系”3 项重要科学研究主题的主题强度在整体时间范围内相对较低，但近 5 年主题强度显著增加，为领域内重要科学研究分支。

2.4 技术研究主题识别与分析

2.4.1 技术研究主题层次聚类结果分析

基于 BERTopic 模型识别国内智慧图书馆领域技术研究主题共 25 个，通过 BERTopic 模型层次聚类功能，展现所有技术研究主题的主题名称与层次聚类情况，见图 5。

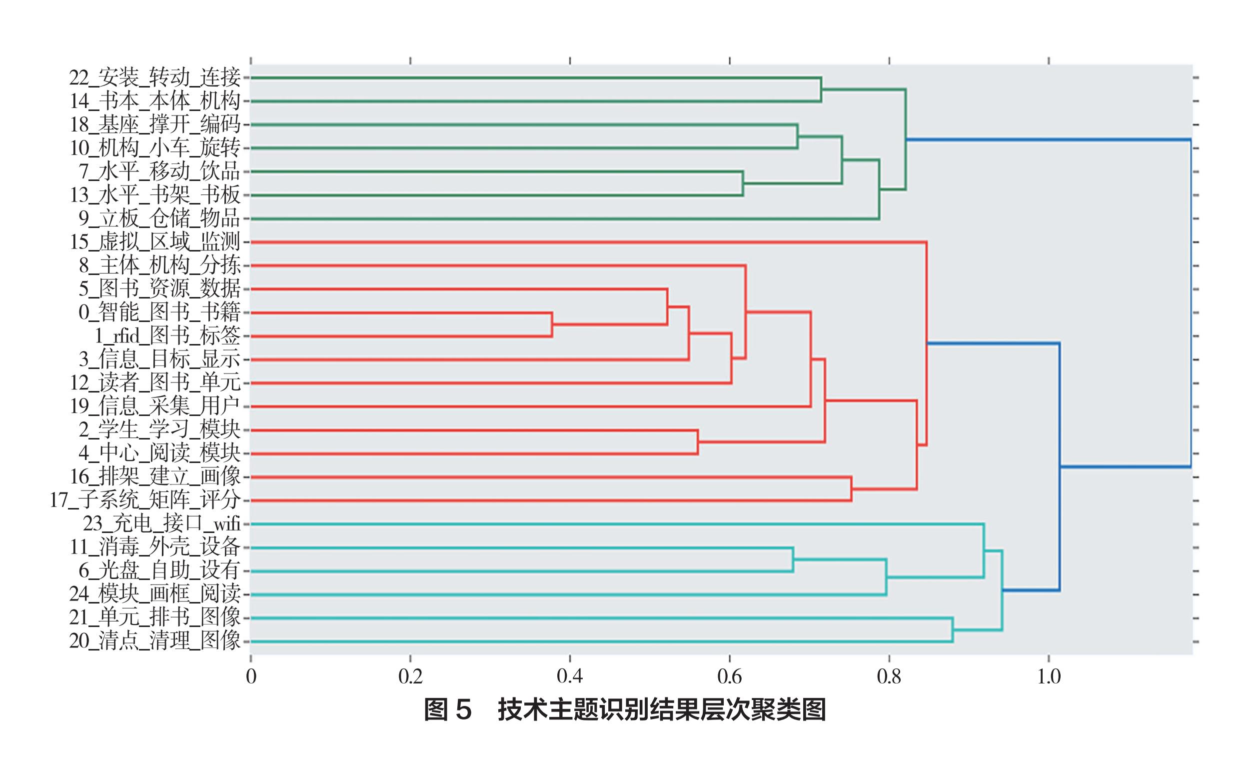


图 5　技术主题识别结果层次聚类图

如图 5 所示，通过层次聚类图谱颜色分布情况可知，国内智慧图书馆技术研究领域大致可以划分为 3 种类别：

（1）深绿色聚类结果：智慧图书馆图书移动与存取技术。该聚类结果下的技术研究主题包括：主题 22（安装 _ 转动 _ 连接）、主题 10（机构 _ 小车 _ 旋转）和主题 13（水平 _ 书架 _ 书板）等。如专利 CN112025668B[28] 公开了一种图书馆内书籍运输车，通过自动识别装置能够有效识别归还书籍的信息，将书籍自动归类放在不同横板上，位于同一个横板上的书籍即为放在同一个书架上的书籍，方便工作人员将书迅速归位。又如专利 CN111015692A[29] 公开一种图书馆自动

存书机械手，该装置结构简单，操作便捷，通过机械传动与机械手的自动化操作，完成市民精准归还图书，极大减少了图书馆员工作量，提高了图书归还效率。

（2）红色聚类结果：智慧图书馆图书智能分拣与个性化阅读技术。该聚类结果下的技术研究主题包括：主题 0（智能 _ 图书 _ 书籍）、主题 1（rfid_ 图书 _ 标签）和主题 19（信息 _ 采集 _ 用户）等。在图书智能分拣技术方面，如专利 CN111797939A[30] 公开了一种基于小波分析的无人图书馆深度学习智能识别系统与方法，实现对图书书脊信息的精准识别，避免依赖传统的 RFID 标签和条形码标签，自动化程度高，无需人工操作。专利 CN107153854B[31] 公开了一种基于超高频 RFID 技术的自动化图书盘点方法，能够对图书标签的相位信号进行双曲线拟合，建立模型，根据模型的系数判断图书在书架的位置。在个性化阅读服务方面，如专利 CN107423413A[32] 公开了一种基于云计算的数字图书馆管理服务系统，通过云计算利用庞大的数据资源和计算能力为数字图书馆提供个性化用户服务，在保证用户隐私情况下，通过行为模式分析用户的浏览、借阅、检索记录等信息，提供更完善的个性化服务。

（3）青绿色聚类结果：智慧图书馆图书智能管理与阅读辅助技术。该聚类结果下的技术研究主题包括：主题 6（光盘 _ 自助 _ 设有）、主题 24（模块 _ 画框 _ 阅读）和主题 20（清点 _ 清理 _ 图像）等。在图书智能管理方面，如专利 CN103295034B[33] 公开了一种基于 DSP 的嵌入式图书乱架清点系统及方法，能够实现图书馆在架图书的自动清点，硬件结构简单，软件代码小，具有功耗低、实时性强、微型化、成本低等特点，可以快速完成图书馆大流通服务的图书清点和盘点工作。在阅读辅助技术方面，如专利 CN107274314A[34] 公开了一种智能化图书馆席位分配系统，具有智能化控制、自主性强的特点，可通过 WIFI 联通手机实时查询座位详情，显著缓解图书馆员工作量，提高图书馆席位的流转率和使用效率。

2.4.2 重要技术研究主题分析

国内智慧图书馆领域主题强度（以主题下专利分布数量表示）大于等于均值的技术研究主题共计 6 项。进一步结合各主题间关键词语义进行二次人工归类与命名，最终获得重要技术研究主题 3 项，合并前后的主题名称、主题强度见表 2。

表 2 国内智慧图书馆领域重要技术研究主题分布情况

合并后重要技术研究主题名称	合并后主题强度（专利数量）	原始技术研究主题名称	主题强度（专利数量）
智慧图书馆信息资源管理系统	338	主题 0- 智能 _ 图书 _ 书籍	291
		主题 5- 图书 _ 资源 _ 数据	47
图书智能标签与检索技术	181	主题 1-rfid_ 图书 _ 标签	123
		主题 3- 信息 _ 目标 _ 显示	58
高校智慧图书馆用户行为分析技术	132	主题 2- 学生 _ 学习 _ 模块	77
		主题 4- 中心 _ 阅读 _ 模块	55

如表 2 所示，“智慧图书馆信息资源管理系统”主题强度最高，表示该技术在智慧图书馆建设中处于核心引领位置，信息资源管理系统是收集、分析、处理、使用各类智慧图书馆数字资源的应用集合体，对智慧图书馆提升资源管理效率、提高服务质量与推动知识传播等具有重要作

用。该主题反映了在大数据和人工智能等数智技术驱动下，智慧图书馆技术领域有关研发者对高效整合、智能分析及精准推送信息资源具有迫切需要与实践导向。

“图书智能标签与检索技术”主题排名第 2，揭示了智慧图书馆技术领域有关研发者对提升智慧图书馆的信息检索效率和图书整理效率方面的重视。图书智能标签与检索技术相辅相成，为图书赋予智能标签，能够实现图书精细化分类与快速定位，极大简化用户检索流程，提高信息获取便捷性。如 RFID 技术可以实现自动化图书借阅、归还、定位和分类等操作，提高了智慧图书馆服务效率。对于图书馆员而言，通过智能标签技术可以实现图书自动分类、上架和盘点，减轻工作负担。对于用户而言，智能检索技术的嵌入，便于用户通过语音等方式表达自身阅读需求，精准快捷检索所需阅读资源。

“高校智慧图书馆用户行为分析技术”主题强度排名第 3，表明智慧图书馆技术领域有关研发者注重利用数智技术对高校图书馆用户行为数据进行收集、处理和分析，揭示用户阅读习惯、兴趣偏好、信息需求等特征。该技术直接关系到高校智慧图书馆服务质量与用户满意度。通过分析高校智慧图书馆用户阅读记录、搜索历史等行为数据，图书馆可为用户推荐符合其兴趣和需求的阅读与学习资源。从科学研究角度看，用户行为分析技术还可以为学术研究提供数据支持，帮助研究者了解学术研究热点和发展态势。

2.4.3　重要技术研究主题演化趋势分析

以时间为横轴、重要技术研究主题的主题强度为竖轴，面向国内智慧图书馆领域 3 项重要技术研究主题进行演化趋势分析，结果见图 6。

如图 6 所示，在 2004—2010 年间，各重要技术研究主题的主题强度经历了小范围“上升—下降”趋势。在 2010—2016 年间，各重要技术研究主题经历了跌宕起伏的动态变化趋势，但彼此间主题强度差距较小。在 2016—2024 年间，“智慧图书馆信息资源管理系统”主题强度同其他两项主题拉开差距，呈现显著引领趋势，于 2018 年达到主题强度高峰，揭示了本主题在智慧图书馆技术体系中的重要性。“图书馆智能标签与检索技术”同“高校智慧图书馆用户行为分析技术”在 2016—2024 年间彼此主题强度差距较小，且整体呈现动态变化趋势，相较于两者 2010—2016 年间主题强度的变化起伏较大，为领域中重要技术研究分支。

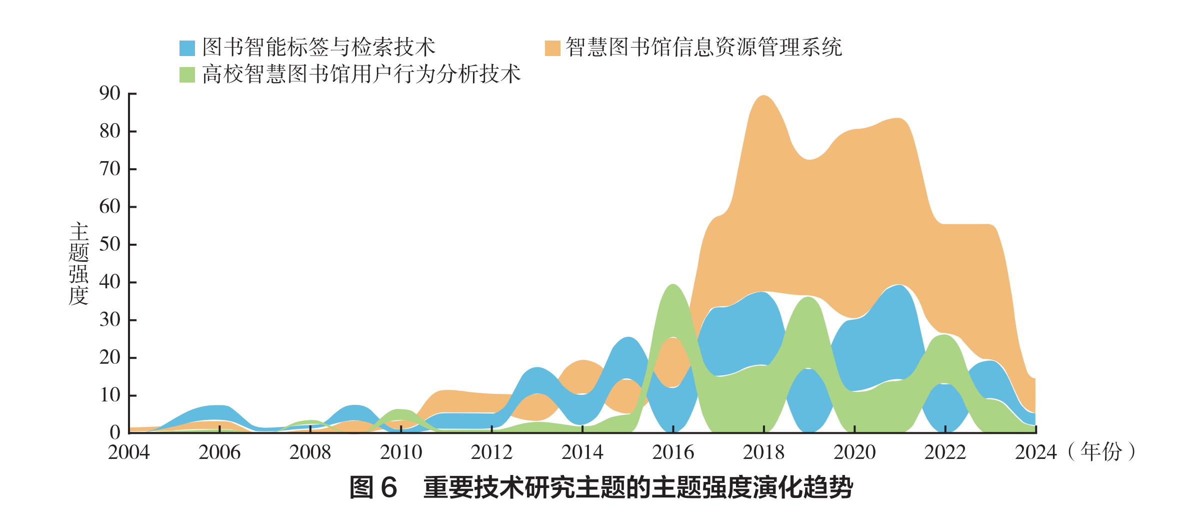


图 6　重要技术研究主题的主题强度演化趋势

3 国内智慧图书馆“科学—技术”主题交互现状分析

将上述 5 项重要科学研究主题所属论文摘要为数据源、3 项重要技术研究主题所属专利摘要为数据源，在进行语料清洗与预处理后，使用 TfidfVectorizer 对语料进行 TF-IDF 向量化表示，进一步两两计算各重要科学研究主题所属论文集合与各重要技术研究主题所属专利集合间的文本平均余弦相似度。当两者间计算结果大于等于均值 0.3151 时，表明“科学—技术”主题间存在交互关系，基于计算结果的“科学—技术”重要研究主题交互情况桑基图见图 7。

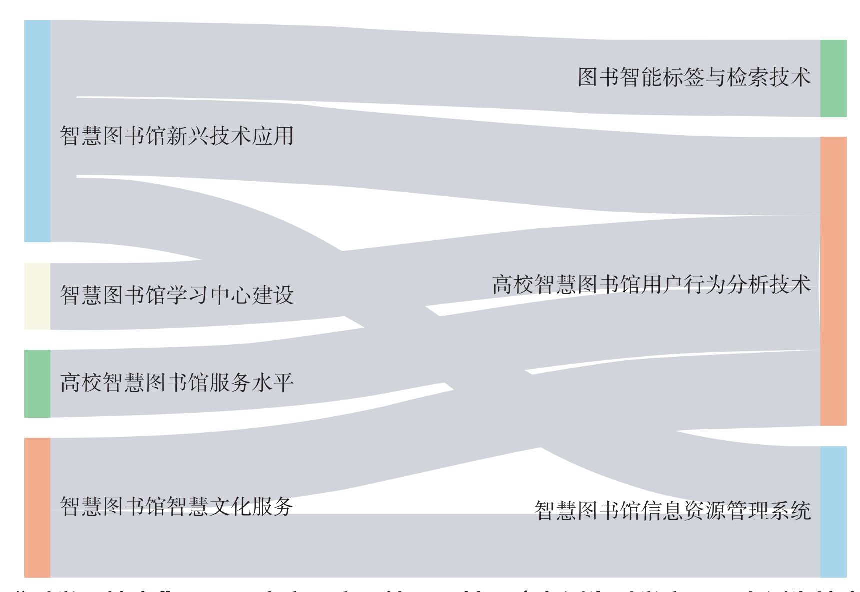


图 7 “科学—技术”重要研究主题交互情况桑基图（左侧为科学主题，右侧为技术主题）

如图 7 所示，科学主题“智慧图书馆新兴技术应用”同各项技术主题交互范围最广，同 3 项技术主题均存在交互关系，表明该科学主题具有较强技术交叉性，在研究内容中广泛借鉴图书智能标签、智能检索与用户行为分析等技术。交互数量排名第 2 的科学主题为“智慧图书馆智慧文化服务”，同技术主题“高校智慧图书馆用户行为分析技术”“智慧图书馆信息资源管理系统”存在交互。而科学主题“公共智慧图书馆文化服务体系”同各项技术主题均不存在交互，由此说明该研究主题下技术支撑相较薄弱，面向技术内容的借鉴程度相对较低，有待进一步深化研究内容的技术支撑。

反之，从技术主题角度看，“高校智慧图书馆用户行为分析技术”主题同各项科学主题交互范围最广，同 4 项科学主题均存在交互关系，表明该技术具备较强知识牵引性，能够为智慧图书馆学习中心建设、高校智慧图书馆服务水平提升、智慧图书馆智慧文化服务提供技术铺垫。交互数量排名第 2 的技术主题为“智慧图书馆信息资源管理系统”，表明信息资源管理系统是国内“智慧图书馆智慧文化服务”与“智慧图书馆新兴技术应用”研究主题下学者重要技术关注方向。“图书智能标签与检索技术”主题同科学主题交互性最弱，仅与科学主题“智慧图书馆新兴技术应用”存在交互，表明该技术在国内智慧图书馆领域科学研究中的探讨较为薄弱，存在较大拓展

研究空间。

从整体上看，国内智慧图书馆领域面向“科学—知识”的交互程度较低，在 15 项交互连接中，仅有 7 项存在交互关系，比例尚未达到一半。同时，除“智慧图书馆新兴技术应用”主题外，其他科学主题存在 0~2 项交互关系，一方面反映了在部分科学研究中缺少对智慧图书馆技术的引荐，技术支撑性薄弱，另一方面也反映出在未来研究中有较大技术引荐、交互提升空间。

4 研究启示

4.1 遵循政策指引与场景化应用导向，促进研究成果从理论向实践转化

由本文科学研究主题识别结果可知，面向有关政策导向下的研究主题较少，同时，面向细粒度场景化智慧图书馆建设主体，如学习中心等的研究主题亦较为匮乏。鉴于此，国内智慧图书馆领域学者可以将自身研究议题，同各项国家、省市级政策中面向智慧图书馆有关的指引内容密切结合。如《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》[1]提出，要深入发展智慧图书馆与智慧旅游，从国家政策层面确立了我国图书馆与旅游事业相融合的智慧化发展战略目标。这启示领域研究者可以深入探索“智慧图书馆 + 智慧文旅”的研究议题，从原则、策略、路径、框架等方面，由点及面开展国家政策引导下且符合国家战略发展需求的智慧图书馆研究议题。领域内学者可以考虑开展面向智慧图书馆不同场景化建设主体有关研究，当前较为聚焦的建设主体为智慧图书馆未来学习中心，以高校师生为主要研究对象。为丰富面向不同场景化建设主体的研究内容，细分实践策略，领域研究者可以探索智慧图书馆“阅读疗法空间”、“红色文献阅览室”、“自助服务区”及“艺术展览室”等建设主体的实践研究，推动研究成果中的理论框架、优化策略、实现路径等向实践转化。

4.2 拓展面向公共智慧图书馆的服务研究，实现“公共馆—高校馆”并重

由本文“科学—技术”主题识别与分析结果可知，当前领域下的研究主体侧重于高校智慧图书馆。如重要科学研究主题中的“高校智慧图书馆服务水平”、重要技术研究主题中的“高校智慧图书馆用户行为分析技术”等，均以高校智慧图书馆为主体，并且有关研究主题的主题强度显著高于面向公共智慧图书馆的有关研究主题。这启示领域内研究者可以拓展面向公共智慧图书馆的服务研究，实现“公共馆—高校馆”并重。相较于高校智慧图书馆，公共智慧图书馆的服务主体更加多元，这启示领域内研究者可以结合各类服务群体的知识结构与阅读需求等，探索公共智慧图书馆特有的服务模式和服务路径，如社区共建共享、公共智慧图书馆联盟等。同时，作为全民阅读推广的重要推动主体之一，公共智慧图书馆如何面向老年人、留守儿童等弱势群体打通“阅读最后一公里”，值得领域内研究者深入探索。

4.3 提升智慧图书馆各类主题研究的可持续性，夯实科学、技术探索深度

由重要科学研究主题与重要技术研究主题演化分析结果可知，两者部分主题在一定时间范围内均存在主题强度跌宕起伏变化趋势，导致面向部分科学与技术主题的研究连续性受到一定程度限制。这启示领域内研究者与技术研发者可以提升智慧图书馆各类主题研究的可持续性，夯实知识、技术的探索深度。面向领域内研究者，可以持续关注“智慧图书馆学习中心建设”“公共智

慧图书馆服务体系”等研究主题，深入探索如何借鉴不同理论、不同技术等提升智慧图书馆学习中心的服务效能，或探索公共智慧图书馆如何开展面向不同社会群体的阅读服务策略等，进一步从横向时间动态发展截面上不断更迭优化前期已有研究，从纵向层面上促进面向上述主题的研究成果向纵深发展，提升知识嵌入深度。对于领域内技术研发者，则可以考虑适当加强对“图书智能标签与检索技术”“高校智慧图书馆用户行为分析技术”主题的持续性研究，迭代优化各类基于物联网、RFID、大数据、云计算、人脸识别等技术的智慧图书馆图书标注、检索、定位和用户行为分析策略等，为智慧图书馆有关科学研究提供新颖性技术信息参考。

4.4　注重引入技术信息支撑，增加“科学—技术”交互程度

由本文“科学—技术”主题交互现状分析结果可知，在 5 项重要科学研究主题与 3 项重要技术研究主题间，仅存在 7 项交互关系，且重要科学研究主题“公共智慧图书馆文化服务体系”同其他重要技术研究主题间均不存在交互关系，这反映了当前智慧图书馆领域科学研究中，面向技术内容的讨论、引入与借鉴不足，技术交融性较为薄弱。领域内研究者可以注重引入技术支撑，提升“科学—技术”交互程度。当前，各类大语言模型技术、人工智能生成内容技术等为当前学术界、技术领域、社会与民生中的关注焦点。为提升研究前瞻性，领域内研究者可以广泛探索各类新兴技术，将其融入智慧图书馆服务内容提供、服务体系优化、图书馆员技能提升、图书馆功能区域建设等的实现原则与微观实践策略，亦可从伦理与隐私角度探讨新兴数智技术应用于智慧图书馆建设时存在的伦理边界与隐私问题等。

【参考文献】

[1] 全国人民代表大会. 中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要［EB/OL］.［2024-08-09］. http://www.npc.gov.cn/npc/c2/kgfb/202103/t20210313_310753.html.

[2] 王飞，徐旭光，缪幽竹. 智慧图书馆文化的研究现状和发展趋势分析［J］. 河南图书馆学刊，2022，42（1）：130-134.

[3] 丁安，褚艳秋，朱朝凤，等. 近十年我国智慧图书馆服务模式研究综述［J］. 图书馆学刊，2019，41（1）：122-126.

[4] 李显志，邵波. 国内智慧图书馆理论研究现状分析与对策［J］. 图书馆杂志，2013，32（8）：12-17.

[5] 裴宏娇，杨志和. 基于专利数据的我国智慧图书馆热点技术发展现状与趋势分析［J］. 科技创新与生产力，2024，4（3）：87-91.

[6] 赵昆. 智慧图书馆相关专利现状与发展趋势［J］. 河南图书馆学刊，2022，42（4）：85-86，94.

[7] 陆伟，李鹏程，张国标，等. 学术文本词汇功能识别——基于BERT向量化表示的关键词自动分类研究［J］. 情报学报，2020，39（12）：1320-1329.

[8] 高凡，徐思佳. 基于BERTopic模型的数字政府治理领域的主题识别与内容分析［J］. 情报理论与实践，2024，47（8）：95-106.

[9] 高春玲，姜莉媛，董天宇. 基于BERTopic模型的老年人健康信息需求主题演化研究——以新浪微博平台为例［J］. 情报科学，2024，42（4）：111-118.

[10] 杨思洛，吴丽娟. 基于BERTopic模型的国外信息资源管理研究进展分析［J］. 情报理论与实践，2024，47（2）：189-197.

[11] 高道斌，李志杰，吴红. 基于引文出版年光谱的国际智慧图书馆领域历史根源文献识别与分析［J］.

文献与数据学报，2024，6（1）：43-57.

［12］Aittola M, Ryhänen T, Ojala T. SmartLibrary-location-aware mobile library service［C］//Human-Computer Interaction with Mobile Devices and Services: 5th International Symposium, Mobile HCI 2003, Udine, Italy, September 2003. Proceedings 5. Springer Berlin Heidelberg, 2003: 411-416.

［13］王静，李新春，尹良伟，等. 基于数字孪生的高校图书馆智慧服务数据治理自适应模式研究［J］. 图书馆，2023（3）：1-7.

［14］陆康，刘慧，杜京容，等. 数据治理——我国智慧图书馆高质量发展新机遇［J］. 图书馆，2022（10）：30-34.

［15］徐小红，薄平平. 基于利他行为心理解读图书馆工作的效用——以江西图书馆为例［J］. 江西理工大学学报，2013，34（2）：66-69.

［16］中国政府网. “十四五”公共服务规划［EB/OL］.［2024-08-09］. https://www.gov.cn/zhengce/zhengceku/2022-01/10/5667482/files/301fe13cf8d54434804a83c6156ac789.pdf.

［17］李睿. “十四五”时期公共图书馆智慧服务推进策略研究［J］. 甘肃科技，2023，39（9）：119-121，125.

［18］杜希林，王宇，车宝晶. 数智赋能公共图书馆服务体系创新的理论思考及实现路径［J］. 图书情报工作，2024，68（7）：3-10.

［19］秦春阳. 高质量发展环境下公共图书馆智慧服务体系优化发展研究［J］. 图书情报导刊，2023，8（5）：1-7.

［20］柳甜，谭海兵，彭峰，等. “双一流”高校图书馆智慧门户设计与实现——以湖南师范大学图书馆为例［J］. 大学图书情报学刊，2023，41（1）：106-112.

［21］杜芃诺. 数据驱动下高校图书馆智慧阅读服务平台构建研究［J］. 图书馆界，2023（1）：48-51，60.

［22］马捷，赵天缘，王思. 高校智慧图书馆功能结构模型构建［J］. 情报科学，2017，35（8）：56-61.

［23］车宝晶，吴瑾. 数智赋能高校图书馆学科服务转型：理论思考、现实依据与实现路径［J］. 图书馆工作与研究，2024（6）：43-51.

［24］陆康. 智慧服务环境下高校图书馆读者隐私保护研究［J］. 图书馆理论与实践，2019（3）：18-24.

［25］闫芳芳. 生成式人工智能赋能智慧图书馆创新的实现策略和风险挑战［J］. 江苏科技信息，2024，41（7）：98-101.

［26］郭军. 智慧环境下图书馆用户个人信息保护研究［J］. 图书馆工作与研究，2023（11）：32-37.

［27］梅振荣，裴丽. 基于区块链的智慧图书馆用户隐私保护［J］. 中华医学图书情报杂志，2019，28（2）：55-59.

［28］同济大学. 一种图书馆取还书机器人系统：CN112025668B［P］. 2020-06-19.

［29］广州基俊机械科技有限公司. 一种图书馆自动存书机械手：CN111015692A［P］. 2020-01-07.

［30］天津中德应用技术大学，天津滨海迅腾科技集团有限公司. 基于小波分析的无人图书馆深度学习智能识别系统与方法：CN111797939A［P］. 2020-07-20.

［31］南京大学. 一种基于超高频RFID技术的自动化图书盘点方法：CN107153854B［P］. 2017-04-24.

［32］华博胜讯信息科技股份有限公司. 基于云计算的数字图书馆管理服务系统：CN107423413A［P］. 2017-07-28.

［33］中南大学. 一种基于DSP的嵌入式图书乱架清点系统及方法：CN103295034B［P］. 2013-04-30.

［34］上海理工大学. 智能化图书馆席位分配系统：CN107274314A［P］. 2017-06-12.

Research on Theme of “Science-Technology” in Domestic Smart Libraries Based on BERTopic

Gao Daobin

(WISE Lab, Institute of Science of Science and S&T Management, Dalian University of Technology, Dalian 116024, China)

Abstract: ［**Purpose/Significance**］Identify and analyze the theme of “science-technology” in domestic smart libraries, and provide information support for scientific research and technological development in this field, and provide reference for the two-way flow of knowledge of science and technology. ［**Method/Process**］Using 5436 Chinese papers from the CNKI journal database and 1416 Chinese patents from the incoPat patent database as data sources, the scientific and technological carriers were characterized respectively. Based on the BERTopic model, the “science-technology” theme was identified, and the clustering results between themes, important research themes, and theme evolution trends were analyzed. On this basis, analyze the interactive relationship between various important scientific research topics and technological research topics, and propose future research implications. ［**Result/Conclusion**］Based on the BERTopic model, scientific research fields are divided into four categories, with five important scientific themes included; divide the field of technological research into three categories, with three important technical themes included; identify a total of seven interactions “science-technology”. It is proposed to follow policy guidelines and scenario based application orientation to promote the transformation of research results from theory to practice; expand the research on services for public smart libraries and achieve equal emphasis on “public library-university library”; enhance the sustainability of various thematic research in smart libraries and solidify the depth of scientific and technological exploration; emphasize the introduction of technological information support and increase the degree of interaction between science and technology, with a total of four research inspirations.

Keywords: Smart library; BERTopic; Scientific theme; Technical theme; Interaction status

（本文责编：魏　进）